Title

Segmentation of Pollen Tube Growth Videos Using Dynamic Bi-Modal Fusion and Seam Carving.

Authors

Tambo, Asongu L
Bhanu, Bir

Publication Date

2016-03-04

DOI

10.1109/tip.2016.2538468

Peer reviewed

Segmentation of Pollen Tube Growth Videos Using Dynamic Bi-Modal Fusion and Seam Carving

Asongu L. Tambo, *Student Member, IEEE*, and Bir Bhanu, *Fellow, IEEE*

***Abstract*— The growth of pollen tubes is of significant interest in plant cell biology, as it provides an understanding of internal cell dynamics that affect observable structural characteristics such as cell diameter, length, and growth rate. However, these parameters can only be measured in experimental videos if the complete shape of the cell is known. The challenge is to accurately obtain the cell boundary in noisy video images. Usually, these measurements are performed by a scientist who manually draws regions-of-interest on the images displayed on a computer screen. In this paper, a new automated technique is presented for boundary detection by fusing fluorescence and brightfield images, and a new efficient method of obtaining the final cell boundary through the process of *Seam Carving* is proposed. This approach takes advantage of the nature of the fusion process and also the shape of the pollen tube to efficiently search for the optimal cell boundary. In video segmentation, the first two frames are used to initialize the segmentation process by creating a search space based on a parametric model of the cell shape. Updates to the search space are performed based on the location of past segmentations and a prediction of the next segmentation. Experimental results show comparable accuracy to a previous method, but significant decrease in processing time. This has the potential for real time applications in pollen tube microscopy.**

***Index Terms*— Growth of pollen tubes, image fusion, seam carving for segmentation, pollen tube model, pollen tube segmentation, pollen tube video segmentation.**

I. Introduction

THE life cycle of the pollen tube from germination to sperm delivery at the ovary is an active area of research in plant biology. Pollen tubes are single cell organisms that are important in the process of sexual reproduction since they are responsible for delivering male sexual material to the ovary. To complete this task, a pollen navigates female tissue by selectively extending a particular part of the cell wall, a process called polar growth [1], [2]. This growth is influenced by many factors (various ions, proteins, pressure), and the interplay between them is another active field of research [1], [3]–[6]. Plant biologists study pollen tubes to determine how these proteins and ions affect external cell characteristics such as cell size, length, growth rate, etc [7]. These external characteristics can only be obtained if the contour of the cell is known: hence the importance of accurate cell segmentation. Furthermore, segmentation is also an important step in automating mathematical growth models that require cell measurements such as length, size and orientation [3], [5], [8], [9].

Image segmentation methods can either be edge-based or intensity-based. Edge-based segmentation methods (e.g., active contours, level sets [10]) seek to minimize an energy cost function by deforming a curve according to image gradients. The optimal curve location lies along image gradients. Other edge-based methods use a predefined criteria to link edges in the image to obtain the best connected edges in the image [11], [12]. Intensity-based segmentation methods (e.g., graph cut [13], [14], region growing [15]) rely on the object having pixel values that are statistically distinct from the background pixels. Region growing adds pixels to the foreground if their value is similar to that of the foreground pixel. Graph cut works by arranging image pixels like the nodes of a graph with a source node and a sink node. The algorithm finds the optimal separation (cut) that sets the pixels as either source or sink pixels. This cut is difficult to obtain in cases where the intensity of the object closely matches the intensity of the background. Active Appearance Models (AAM) learn a series of shapes by combining both gradient and texture information [16], [17]. These models usually require a training phase. A new shape is detected as a linear combination of the previously learnt shapes. Given the potential for tremendous variation in cell shapes between experiments, such a model will require constant retraining.

The idea of fusing information from multiple agents to solve a problem is not new in image processing [18]–[24]. Fusion can be done at an intermediate step or at the final step of the process. Han and Bhanu [18], fuse color and infrared video to enhance human detection in color images. Images from infrared sensors have better silhouettes since they are not strongly affected by lighting conditions. These silhouettes are used to improve image registration and the detection of humans in color images. A similar fusion strategy is used by Zou and Bhanu [20] for tracking humans using audio and video data. Audio and video features are fused at the feature-level to improve the performance of human tracking. Parvin et al. [21] use an iterative voting technique to detect cell centers. The voting is done by projecting image edges inwards using progressively focused kernels. This method distinguishes individual cells even when they are overlapping. This method is geared towards detecting object centroids at very small scales. While these methods are sufficient for detection,

Manuscript received August 21, 2015; revised December 18, 2015; accepted February 17, 2016. Date of publication March 4, 2016; date of current version March 18, 2016. This work was supported by the National Science Foundation under Grant DGE 0903667 and Grant IIS 0905671. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Oleg V. Michailovich. *(Corresponding author: Asongu L. Tambo.)*

The authors are with the Department of Electrical and Computer Engineering, University of California at Riverside, Riverside, CA 92521 USA (e-mail: atamb001@ucr.edu; bhanu@cris.ucr.edu).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the author.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2538468

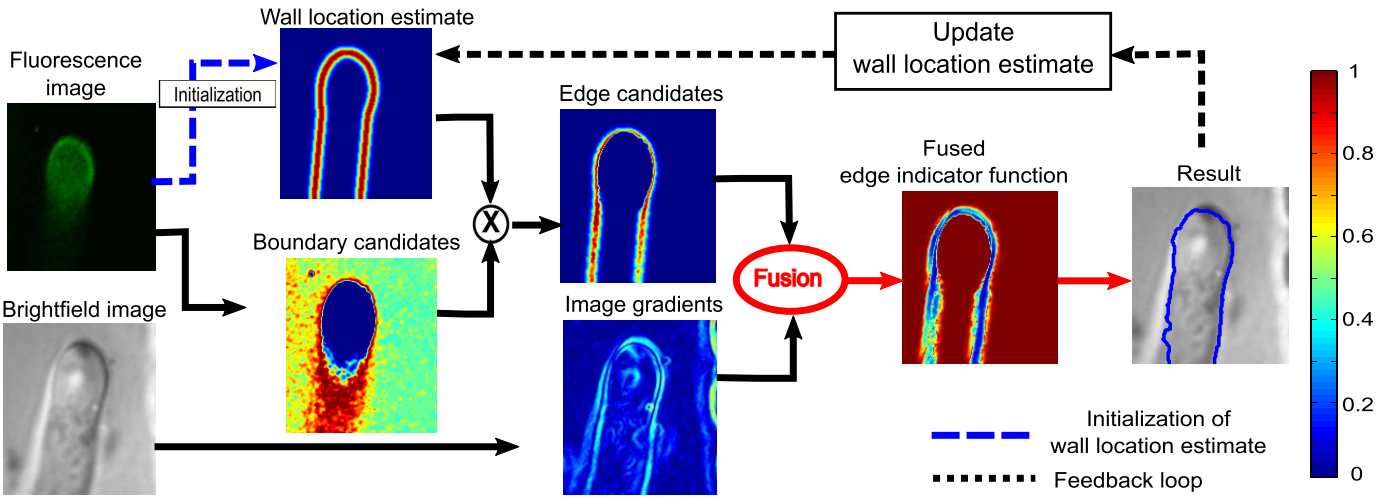

Fig. 1. Block diagram of the image analysis method showing sample results after each stage of the process.

segmentation is a challenge due to the complexity of the images (complex human motion; multiple overlapping cells at a small scale). Tarabalka et al. [24] propose a graph cut method for segmenting videos with spatio-temporal constraints. This method is extended to multi-modal images with the constraint that the segmentation of one modality should be included in that of the other. The segmentation works on groups of images while enforcing continuity in the pixel labeling. The above methods also do not require any human interaction (agents) to validate initial detection estimates.

Warfield et al. [19] propose a statistical method for fusing results from many *agents* to a single segmentation by maximizing the likelihood of the final segmentation given the performance criteria of the agents. This method was later improved by Liu et al. [23] to include image intensity values. This approach is similar to that proposed in Asman and Landman [22] where many agents ($N \geq 8$) are used to statistically arrive at a single segmentation. While these methods would provide accurate segmentations, they require input from multiple agents who understand the presented image. This knowledge greatly improves the quality of the initial segmentations as agents can tell which parts of the image are object/background.

In this paper, we present a method for obtaining a final segmentation by fusing images from two modalities. This model is an extension of our previous work [25]. The key contributions of the method include: 1) a more accurate model for describing the shape of a pollen tube in 2D. Model equations are parametric in nature and provide a means of determining the validity of the initial segmentation. 2) A faster method for obtaining the cell boundary based on energy minimization within the restricted search space. This is based on ideas from *seam carving* that have been used for media retargeting [26]. 3) Temporal updates to the search space based on previous segmentations and motion vectors between the last two segmentations. 4) The two sensors used in this study are fluorescence and brightfield image modalities. Fluorescence images provide localization information, from which the location of the cell wall in the image can be estimated. A fused edge indicator function (FEIF) is computed by combining estimates from the fluorescence image and gradients from brightfield images. This combination serves to amplify desired edges (those in the search space) in the brightfield image while simultaneously attenuating undesired edges. The optimal boundary is located within the search space using the proposed method. This particular way of localizing optimal boundary within two boundary estimates is novel in the image processing field. 5) Experimental results are presented using multi-modal videos for diverse pollen tube growth scenarios.

In the following, we provide a theoretical description of the key aspects of our method in Section II, the results and comparisons of our method with other segmentation algorithms are given in Section III, and the conclusions of this paper are presented in Section IV.

II. Theoretical Development

In this section, we present the theoretical development of our method. Figure 1 is a block diagram highlighting the key steps of our method. Fluorescence images used in this study show the localization of the protein RIC3 [2]: bright pixels indicate high concentrations of RIC3. An initial estimate of the cell is obtained by segmenting the fluorescence image using intensity-based methods (e.g. graph cut [14]). From this segmentation, we determine the likely location of boarder/boundary pixels. For a given object shape, we require a means of determining whether or not the shape is a valid pollen tube. Since the initial segmentation is not performed by an *expert* (a human who can glean the cell shape even when the outline is not complete), we need a way of validating the initial segmentation. If the object does not match the description of a pollen tube, then we fill-in the missing parts of the object by performing shape estimation. Combining this shape estimate with the likelihood of border pixels narrows the search space for the cell wall. Finally, we fuse the boundary estimate with image gradients from the corresponding brightfield image. This serves to attenuate undesired edges. The cell boundary is determined as the path of least energy from one end of

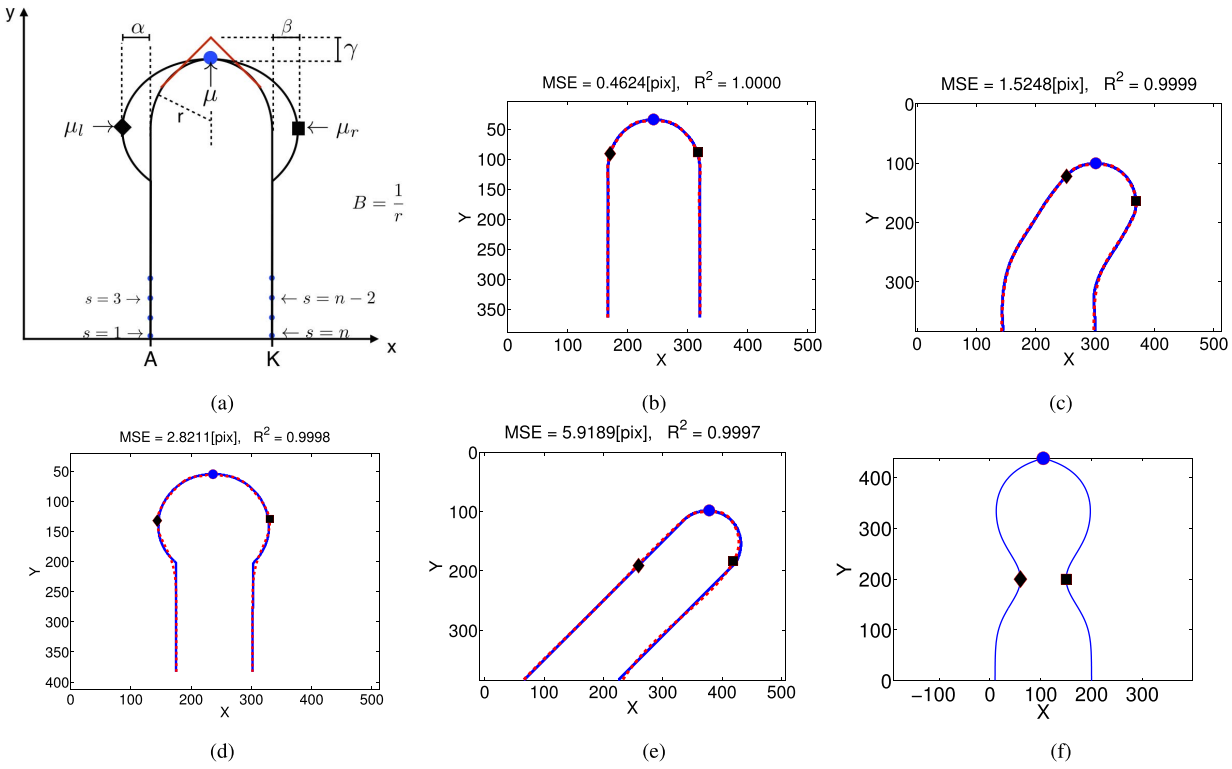

Fig. 2. (a) Artist's rendition of the overlap between two types of pollen tube shapes (straight and bulging) indicating the parameters of equation (1). (b-e) Results of fitting equation (1) (red; dashed) to sample growth scenarios (blue; solid). The diamond (♦) and square (■) markers indicate the centers of the left and right deformation Gaussians respectively. The circle (•) marker indicates the vertex (μ) of the shape. (f) A simulated shape showing changing tube width that causes a pinching-effect. (a) Sample pollen tube shapes. (b) Straight. (c) Turning. (d) Bulging. (e) Slanted. (f) Simulated shape.

the search space to the other. Finally, the search space is updated by the locations of the last N-segmentations through a voting matrix and a prediction of the location of the next segmentation. This prediction is based on the motion vectors between the previous two results.

A. Geometry of a Pollen Tube

A pollen tube is a 3D, cylindrical tube with a spherical dome. In most theoretical developments, it is assumed to be symmetrical in the axial (growth) direction and hence approximated as a 2D cylinder with a circular tip. In [25] we proposed two functions that can be used to represent a straight pollen tube with a vertical orientation in the image plane.

In this paper, we modify those functions to accommodate certain characteristics that pollen tubes exhibit during growth such as turning and bulging (see Figure 2c, 2d). Let $g(s) = \{x(s), y(s)\}$ be the points along the contour of a 2D pollen tube. The equations representing the variation of these points from one end of the tube to the other is given by:

$$x(s) = A + \frac{K - A}{\left(1 + \zeta e^{-B(s-\mu)}\right)^{1/\zeta}} + \alpha e^{-0.5\left(\frac{s-\mu_l}{\sigma_l}\right)^2} + \beta e^{-0.5\left(\frac{s-\mu_r}{\sigma_r}\right)^2} \quad (1a)$$

$$y(s) = m(s - \mu)sign(\mu - s) + m\mu - \gamma\, e^{-0.5\left(\frac{s-\mu}{\sigma_y}\right)^2} \quad (1b)$$

where $[A, K, B, \alpha, \beta, \zeta, \mu, (\mu, \sigma)_l, (\mu, \sigma)_r]$ are 11 parameters that control the dynamics of the x-coordinate, and $[m, \mu, \gamma, \sigma_y]$ are the parameters that control the dynamics of the y-coordinate. μ is common to both functions.

In equation (1a), the first two terms constitute a generalized logistic function which dictates the overall dynamics of the x-coordinate. The third and fourth terms are most influential in deforming the shape from the expected logistic behavior on the left and right sides respectively. During vertical growth with minimal bulging, both α and β are much smaller than A or K and the parameters (A, K) represent the left and right walls of the tube respectively. μ is the index of the coordinate of the tube vertex, and ζ depends on initial conditions. In the case of turning, tip bulging, and other growth behavior, the interpretation of the parameters of equation (1a) is no longer straight forward.

Equation (1b) attributes the behavior of the y-coordinate of the shape to a combination of a triangle function and a Gaussian function. The Triangle function (first two terms) dictates the overall shape of the sequence. The last term (Gaussian term) flattens out the middle of the shape and contributes to the curved nature of the tip. m is the gradient of the triangle function and is either -1 or 1. a is the index of the vertex. This index is more reliable than μ from equation (1a) because the y-coordinate is robust to turning or bulging of the cell shape. This is due to the assumption that all pollen tubes are growing from the bottom of the image to the top.

TABLE I

TABLE DESCRIBING THE PARAMETERS USED IN EQUATIONS (1) AND THEIR LIMITS. *col* REFERS TO THE WIDTH OF THE IMAGE IN PIXELS AND L REFERS TO THE NUMBER OF POINTS IN THE SHAPE CONTOUR

Parameter	Description	Limits
A	Position of the left wall of the pollen tube (when there is no slanting)	$[0, \quad col]$
K	Position of the left wall of the pollen tube (when there is no slanting)	$[0, \quad col]$
B	Radius of curvature of a circle fitted to the tip region	$[0, \quad 1]$
ζ	Depends on the initial value ($s = 0$) of the sequence	$\frac{1}{4}[-L, \quad L]$
μ	Index of the location of the vertex in the sequence	$\frac{1}{4}[L, \quad 3L]$
α	Amount of deviation of the curve along the left wall. Value is < 0 if bulging and > 0 if indentation	$\frac{1}{2}[-L, \quad L]$
β	Amount of deviation of the curve along the right wall. Value is < 0 if indentation and > 0 if bulging	$\frac{1}{2}[-L, \quad L]$
μ_l	Turning point of the shape deviation in the left wall	$\frac{1}{2}[-L, \quad L]$
σ_l	33% of the width of the shape deviation in the left wall	$\frac{1}{2}[0, \quad L]$
μ_r	Turning point of the shape deviation in the right wall	$[\frac{L}{2}, \quad L]$
σ_r	33% of the width of the shape deviation in the right wall	$\frac{1}{2}[0, \quad L]$
m	Gradient of the first half of the sequence	-1 or 1
γ	Height of the flattening Gaussian function	$\frac{1}{2}[-L, \quad L]$
σ_y	Standard deviation of the flattening Gaussian	$[0, \quad \frac{L}{2}]$

TABLE II

TABLE OF PARAMETER VALUES FROM FITTING EQUATION (1) TO THE SHAPES SHOWN IN FIGURE 2. IN ALL INSTANCES, $m = -1$

Figure	Parameters												
#	A	K	B	μ	ζ	α	μ_l	σ_l	β	μ_r	σ_r	γ	σ_y
2(b)	166.64	320.45	0.03	360.01	0.95	-10.31	273.64	33.72	11.40	446.49	34.23	-26.45	24.38
2(c)	142.01	297.98	0.02	311.99	-0.01	88.54	265.51	80.30	79.89	407.12	64.91	-25.15	23.52
2(d)	175.54	301.87	0.03	368.50	0.81	-34.14	252.43	40.66	32.83	484.42	41.92	-35.13	31.94
2(e)	0.00	189.15	0.04	309.04	5.43	136.02	160.63	115.45	227.78	390.19	125.33	-20.97	20.14
2(f)	10	200	0.08	350	1	50	175	50	-50	525	50	-34.5	31.7

γ is the height of the Gaussian function that flattens out the middle of the sequence and σ_y is the standard deviation. Table I summarizes the parameters of equation (1) and lists their bounds.

Figure 2(a) shows an artist's rendition of two overlapping pollen tube shapes and marks the locations of the parameters used in equation (1). Figure 2 also shows some sample growth scenarios of pollen tubes (blue solid lines) and the result of fitting equation (1) to the extracted shape shown in blue. The Mean Squared Error (MSE) of the function is also indicated. Figure 2(b) shows a vertical pollen tube that is the model situation. The left and right Gaussians (diamond and square markers) are located at the start and end of the curved tip and slightly modify the general logistic shape. In Figure 2(c), the left Gaussian is located where the vertex should have been and 'flattens' the expected turning. The right Gaussian compensates for the overshoot of the x-values beyond the expected maximum value of x, which is K. In Figure 2(d), the left and right Gaussians are centered at the middle of the bulging in their respective walls. Finally, Figure 2(e) shows a slanting pollen tube. The left Gaussian is centered at a negative index, indicating that there is no vertical growth captured in the left wall. The right Gaussian captures the tip but has a large standard deviation (σ) to compensate for the overshoot in x-values. Figure 2(b) is a simulated shape that shows what happens during imaging when the center of the cell moves above/bellow the focal plane of the imaging system, and then returns to this plane. This gives the cell a pinched-shape. Table II shows the parameters obtained from the shapes shown in Figures 2b to 2e.

B. Shape Initialization and Boundary Candidates

The fluorescence images used in this study show the localization of RIC3. Regions of the cell containing this protein emit light of a specific wavelength when excited with light of a different wavelength. Since RIC3 is only found inside the cell (and not expected to be in the media surrounding the cell), this modality provides excellent localization for the entire cell. However, there are certain problems that arise from fluorescence microscopy for cell detection: 1) the protein of interest is not uniformly distributed within the cell; 2) certain regions of the cell are out of focus when the cell tip grows out of the plane of focus; 3) photobleaching: prolonged excitation of the specimen causes a reduction in the intensity of the emitted light. These cases produce dark regions in the cell that lead to incomplete detections of the cell shape.

An initial estimate of the location for the cell wall boundary is obtained using the fluorescence image. Due to the influence of the microscope point spread function, the likelihood of classifying a pixel as an object (pollen tube) pixel follows a Gaussian distribution, centered on the expected brightness of the detected region $\left(\mu_{body}\right)$. Consequently, the likelihood

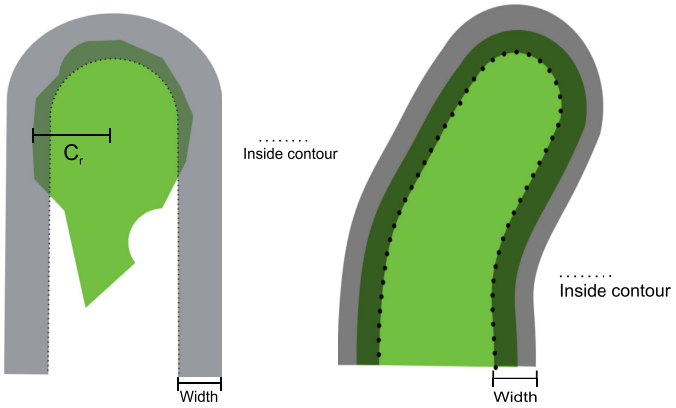

Fig. 3. Artist's rendition showing an overlay of the search space on the initial segmentation. $width = \omega$. The region in Green signifies the fluorescent region of the pollen tube. (a) The initial segmentation is incomplete, so shape estimation is used to fill in the missing parts of the cell. (b) The initial segmentation is complete, so the search space is the group of pixels around the initial contour that are within a distance of $\frac{\omega}{2}$ pixels from the contour. (a) Incomplete initial segmentation. (b) Complete initial segmentation.

of a pixel belonging to the object boundary increases as the pixel intensity approaches the lowest cut-off value of the segmented pixels belonging to the pollen tube. We propose that the intensities of the pixels along the cell wall should satisfy the following distribution:

$$P^{I}_{cw}(i \in border) = \frac{1}{\sqrt{2\pi}\sigma_{body}} e^{-0.5\left(\frac{I(i)-\mu_{wall}}{\sigma_{body}}\right)^2} \tag{2}$$

where μ_{body} and σ_{body} are the mean and standard deviation of the pixel intensities in cytoplasm region (as detected by the initial segmentation), and $\mu_{wall} = min(I_{body}) - \sigma_{body}$. This limit focuses the search of the cell wall region to those pixels whose intensities are close to the threshold of the initial segmentation.

To initialize the search space for the cell wall, we consider two possibilities: incomplete vs. complete initial segmentations. These two are highlighted in Figure 3. The classification of the initial segmentation is important in deciding whether shape estimation is needed. The initial segmentation is classified using the following criterion:

$$C_{init} = \begin{cases} \text{incomplete} & R^2_{equation(1)} < 0.90 \\ \text{complete} & \text{Otherwise} \end{cases} \tag{3}$$

where $C_{init} = \{(x, y)\}$ is the contour of the initial segmentation and R^2 is the r-squared value derived from fitting equation (1) to C_{init}. $R^2_{(.)} = 1 - \frac{\sum_s\{(y(s)-\hat{y}(s))^2+(x(s)-\hat{x}(s))^2)\}}{\sum_s\{(y(s)-\bar{y})^2+(x(s)-\bar{x})^2\}}$. $(\hat{x}, \hat{y})$ are the modeled contour points from equation (1) and $(\bar{x}, \bar{y})$ are the mean of the data points (x, y) respectively.

If the initial segmentation does not satisfy equation (1) (as in Figure 3(a)), then there are parts of the cell that were not captured by segmenting the fluorescence image. This region is most likely the cylindrical part of the cell, since it is this region that suffers from the issues mentioned above. Since we have equations to describe what a complete pollen tube should look like, we can estimate the missing parts of the cell and complete its geometry. Let the points $\rho = \{\rho_x, \rho_y\}$ represent the line of symmetry of the detected shape after the initial segmentation. The best predictor of the shape of the cylinder is either a linear function (for a straight pollen tube) or an exponential function (for a curved tube):

$$f_{sym}(\rho, \mathbf{w}) = \begin{cases} w_1\rho_x + w_2\rho_y + w_3 = 0 & \text{if } R^2_{lin} > R^2_{exp} \\ w_1 e^{w_2\rho_x} + w_3 e^{w_4\rho_x} - \rho_y = 0 & \text{Otherwise} \end{cases} \tag{4}$$

where $R^2_{(.)}$ is the R-squared value obtained by fitting either the linear or exponential function to the line of symmetry of the cell shape. Using equation (4), the estimated optimal location of the cell wall is obtained using:

$$P^{IC}_{shape}(i) = \frac{1}{\sqrt{2\pi}\sigma_m} e^{-0.5\left(\frac{d(i, f_{sym})-C_r}{\sigma_m}\right)^2} \tag{5}$$

where $d(i, f_{sym})$ is the distance of pixel i from the line of symmetry f_{sym}. C_r is the radius of the cylindrical part of the cell. $\sigma_m = \frac{\omega}{3}$ where ω is the radius of the search space within which the cell boundary is expected to be found. This is a user-defined parameter.

If the initial segmentation is complete according to equation (1), then the search space is a band of pixels around the object contour that are within a distance of ω pixels from the object contour. The search space follows the contour of the initial segmentation as shown in Figure 3b and is described by the equation:

$$P^{C}_{shape}(i) = \frac{1}{\sqrt{2\pi}\sigma_m} e^{-0.5\left(\frac{d(i, C_{init})}{\sigma_m}\right)^2} \tag{6}$$

where C_{init} is the contour of the initial segmentation and $d(i, C_{init})$ is the distance of point i to the contour C_{init}. For subsequent images in the video sequence, P_{shape} will be updated based on the motion of individual contour points between consecutive segmentations. This process is described in Section II-E.

Combining equations (2), (5) and (6) gives:

$$P_{CW} = P^{I}_{cw} \times P_{shape}, \text{ where } P_{shape} = \begin{cases} P^{C}_{shape} & \text{If complete} \\ P^{IC}_{shape} & \text{Otherwise} \end{cases} \tag{7}$$

which indicates the most likely location for the cell wall in the image. Shape estimation (equation (5)) is beneficial to the final result when the initial segmentation is incomplete.

C. Fused Edge Indicator Function (FEIF)

A *fused edge indicator function* is obtained by combining the result of equation (7) with the gradient-magnitude of the brightfield image. Brightfield images are formed by capturing the intensity of transmitted light through a specimen. The amount of light transmitted through a particular region depends on the refractive index of that region. The refractive index (η) of a material is the ratio of the speed of light in vacuum (c) to the speed of light in the material (v): $\eta = c/v$. Regions of the sample with relatively high refractive indices (cell wall, mitochondria, golgi body, nucleus, etc)

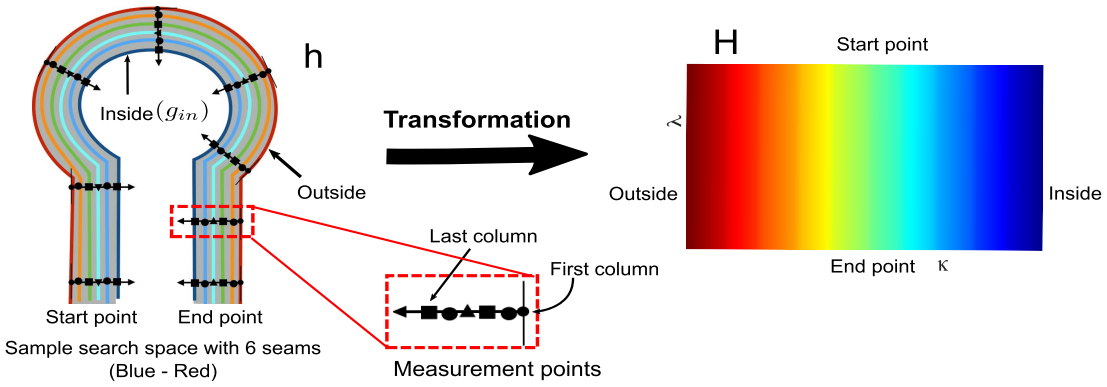

Fig. 4. Transformation of the search space around the contour of an object into a 2D grid for easy searching. [Left] Search space (gray region) for the final segmentation showing 6 seams χ_j (blue - red) out of many. The points belonging to the same seam are equidistant from the *inside* curve (g_{in}). Measurement points lie in a normal direction between the *inside* and *outside* curves. [Right] Search space after transformation. Each column represents a seam and each row is a set of measurements from the *inside* of the space to the *outside*. The first row has measurements from the *start point* and the last column has measurements from the *end point*. Each colored line represents one of many potential segmentations.

will absorb more energy than regions with lower refractive indices (cytosol, media). This causes pixels associated with the former regions to be darker than their surroundings in the image. Since these cytosolic bodies are not necessarily touching, the cell body becomes crowded with edges due to changes in the amount of transmitted light in adjacent patches. At the cell wall/boundary, we would expect to see a contiguous region of dark pixels surrounding the object. However, the orientation of the light source affects the occurrence of this contiguous region. In some images, the cell has a region of lower-intensity pixels on one side, and a region of higher-intensity pixels on the other side. Both regions have high gradients, but one is a minima while the other is a maxima region in intensity. As such, thresholding the image for either one of these regions will not yield a complete cell boundary.

Utilizing the localization estimate from equation (7) serves two purposes: (1) eliminate the false edges found within the cell due to changes in the refractive indices of adjacent patches; (2) emphasize/reinforce the edges that are along the cell boundary. We propose the following fused edge indicator function based on the model suggested by [27]:

$$h = \frac{1}{1 + (P_{CW})_n |\nabla (G_\sigma * I)|} \tag{8}$$

where $*$ is the convolution operation, G_σ is a Gaussian smoothing function, σ is its standard deviation, and $(\bullet)_n$ is a normalization function such that $(\bullet)_n \in [0 \;\; 1]$. This normalization ensures that only the gradients in the undesired regions are attenuated.

D. Detecting The Cell Boundary

The final object segmentation is obtained by finding the line of least energy in the search space starting from one end of the cylindrical part to the other end. In [25] we used level set evolution [10] to find this contour. Parameters for curve evolution were chosen to ensure that the function terminated within the search space. Here we present a new and faster method of obtaining the cell contour using the well known seam carving method [26], which has been widely used for image retargeting by removing (or adding) an optimal 1D vertical or horizontal path of pixels in an image. The authors add that this solution can be obtained using either dynamic programming or graph cut. We shall present a brief summary of the seam carving method and show how the segmentation task restructured as a seam carving problem.

Consider an image I of size $(p \times q)$. A vertical path/seam $(\theta_j \in \Theta = \{\theta_1, \ldots, \theta_n\})$ through I is a collection of pixel locations of length p from the top of the image to the bottom of the image with only one pixel from each row [26], i.e. $\theta_i = \{(x(j), j)_{j=1}^{p} : |x(j) - x(j+1)| \leq 1\}$. To reduce the number of columns in image I by 1, we have to remove a vertical seam/path. The targeted seam (θ^*) is the seam that has the lowest energy, i.e.

$$\theta^* = arg \min_{\Theta} E(\theta) \tag{9}$$

where E is an appropriate energy function.

A global solution to equation (9) can be obtained using the following scheme: starting from the second row of I and for each pixel, check the 8-connected neighbors in the previous row and add the value of the smallest neighbor to the current pixel value. Repeat this procedure up to and including the last row. The smallest value of the last row denotes the seam with the smallest energy. Back-track from this location to the first row, noting the visited locations. The list of locations constitutes the global solution of equation (9).

For our segmentation problem, we are looking for a path of pixels from specified start/end positions that belong to the contour of the cell. This path/contour will have the smallest cost among all available contours. This problem can be recast as a seam carving problem with the following adaptations: 1) since the cell always intersects with the edges of the field-of-view, we use these locations to define start/end positions for the object contour. These positions have the same length (determined by the width of the initial search space) and correspond to the top/bottom of the path needed in seam carving (see Figure 4); 2) we use a transformation $\Psi : (i, j) \rightarrow (\kappa, \lambda)$ to convert the search space to a homogenous region where the

seam carving method can be applied. Ψ is defined as:

$$\Psi : (i, j) \rightarrow (\kappa, \lambda) : \begin{cases} \kappa = d_{\perp}\big((i, j), g_{in}\big) \\ \lambda = d\big((i, j), \text{start point}\big) \end{cases} \tag{10}$$

where κ is the perpendicular distance ($d_{\perp}$) of point (i, j) from the curve g_{in}. This signifies traversing the search space from the *inside* contour to the *outside* contour in a normal direction and is represented by the different colors and measurement symbols in Figure 4. λ denotes the distance of a point from the starting point of the curve that this point belongs to, i.e. λ signifies traversing the search space from *start point* to *end point*. $H(\kappa, \lambda)$ holds the pixel value from position (i, j) of the FEIF (h).

Similar to seam carving, let $\chi_j \in \mathbf{X} = \{\chi_1, \ldots, \chi_N\}$ be the set of vertical seams of length λ in the transformed search space H. The seam with least energy is found by minimizing the following equation (similar to equation (9)):

$$\chi^* = arg \min_{\mathbf{X}} \sum_{j=1}^{N} H(\chi_j) \tag{11}$$

We can now use dynamic programming to solve equation (11) for the optimal contour in H.

E. Updating the Search Space (Wall Location Estimates)

The above segmentation procedure holds for groups of images where each group contains one fluorescence and one brightfield image. In Section II-B the initialization of P_{shape} forms a uniform region around the observed initial segmentation (see Figure 3). While frames from a video sequence can be segmented independent of each other according to the above procedure, there is some additional benefit of knowing that the images are from the same video sequence, i.e. using past segmentations to aid in future segmentations. Over time, the search space should incorporate previous data but also anticipate future growth trends based on immediately observed data. In this section, we discuss methods of updating the search space during video segmentation.

The search space is updated based on an accumulation of previous segmentations and the motion vectors between the two most recent segmentations. This combination allows the space to emphasize the locations of previous segmentations while also adjusting for the next segmentation. This is expressed as a weighted sum:

$$P_{shape} = (1 - \omega) P_{next} + \omega P_{emph} \tag{12}$$

where $\omega \in [0 \;\; 1]$ is a weighting term. In our experiments, we set $\omega = 0.9$ which emphasizes the previously registered segmentation locations (including the current segmentation) as the candidates for the next segmentation. P_{emph} contains a history of the previous segmentations. This history is captured in a voting space, V, which accumulates the votes that each pixel location has received as belonging to the cell contour. After each segmentation, the votes of the pixel locations that belong to the contour are incremented by 1. V is used in the search space as:

$$P_{emph} = P_{prev}\left((G_{\sigma} * V)_n + 0.5\right) \tag{13}$$

Algorithm 1 Pseudo-Code for Video Segmentation

Require: F: Fluorescence video
Require: B: Brightfield video
Require: ω: Initial radius of search space
while video has not ended **do**
  $[f_i, b_i] \leftarrow$ next frames from F and B
  if frame number ≤ 2 **then**
    Initialize search space (P_{shape}): equations (1)$\rightarrow$(6)
  end if
  Compute boundary candidates from f_i: equation (2)
  Amplify/attenuate boundary candidates: equation (7)
  Fuse image gradients from b_i and boundary candidates: equation (8)
  Transform search space + interpolation: equation (10)
  Find object contour by traversing H: equation (11)
  if frame number ≥ 2 **then**
    Update search space: equations (12) $\rightarrow$ (14)
  end if
end while

where G_{σ} is a smoothing Gaussian, $(\bullet)_n$ is a normalization such that $(\bullet)_n \in [0 \;\; 1]$. The above equation amplifies the locations that have received more than half the votes at each iteration and attenuates those locations that have received less than half the votes. This strategy eventually eliminates less visited locations.

To anticipate the next segmentation region, the search space uses the motion vectors between the previous two segmentations (χ^{t-1} and χ^t) [28]. The intuition for this development is that if a location has moved a distance of D pixels, then we expect its next location to be centered D pixels from the current location and moving in the same direction. The motion vectors are computed in the normal direction to χ^{t-1}.

$$P_{next}(i) = e^{-0.5\left(\frac{d(i,\chi^t) - |\vec{D}(j)|}{\sigma_m}\right)^2} \tag{14}$$

where $d(i, \chi^t)$ is the distance between point i and the curve χ^t, $\vec{D} \in \{\text{motion vectors along } \chi^t\}$ and j is the index of the point on χ^t that is closest to point i. This function creates a Gaussian search space around the expected segmentation of the next image: a predicted segmentation region. Combining the above equation with equation (13) updates the search space according to equation (12). This updated search space is used in segmenting the next image. Algorithm 1 outlines the pseudo-code of the video segmentation method.

III. Experimental Results

A. Datasets

There are seven bi-modal datasets of growing pollen tube videos used in this study. These videos are used to perform 10 experiments to demonstrate the strength of our method in obtaining new segmentations that are either close to (Exp. 1-7) or far from (Exp. 8-10) the previous segmentations. Table III gives some properties of each video and the analysis sampling rate for each experiment. The acquisition rate column indicates the microscope waiting time between consecutive pairs of fluorescence and brightfield images.

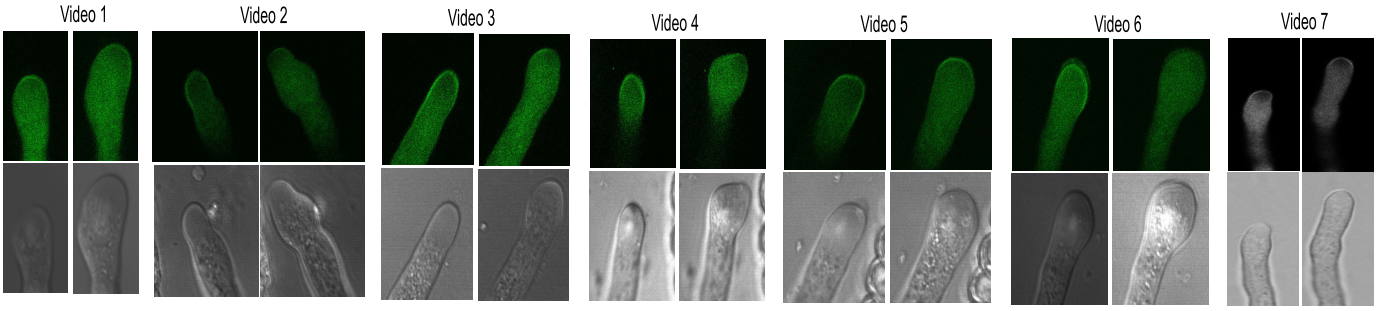

Fig. 5. First and last frames from the videos used in this study. The top row shows fluorescence images and the bottom row shows brightfield images.

TABLE III
PROPERTIES OF THE EXPERIMENTAL VIDEOS

Exp. #	Acquisition rate $[sec/frame]$	Duration $[sec]$	# of frames	Sampling time $[sec]$	# Analyzed frames
1	3.00	501.00	167	3.00	167
2	3.00	459.00	153	3.00	153
3	3.00	228.00	76	3.00	76
4	3.00	228.00	89	3.00	89
5	3.00	228.00	184	3.00	184
6	3.00	228.00	98	3.00	98
7	0.14	58.80	424	0.42	141
8	—— Same as Exp # 1 ——			30.00	16
9	—— Same as Exp # 2 ——			30.00	15
10	—— Same as Exp # 7 ——			1.43	42

The sampling time indicates the time between analyzed images (not processing time). Each dataset consists of a fluorescence video and a brightfield video of the same pollen tube. Each dataset is made during an experiment by taking successive fluorescence and brightfield images at each imaging time point. These images are then made into videos, one for each modality and the two videos make up the dataset. The images are of size 512 × 512 pixels (Exp. 1 - 6) and 512 × 200 pixels (Exp. 7) with a pixel scale of 0.0582 $\mu m/pixel$ and were obtained using a Leica confocal microscope. In the analysis, image dimensions greater than 256 were resized to 256. i.e. (512×512) → (256×256) and (512×200) → (256×200).

As listed in Section II-B, fluorescence microscopy suffers from certain problems that give rise to an incomplete image of the pollen tube (Figure 5). These include the absence of fluorescent markers in certain regions of the cytoplasm as well as loss of intensity due to over-exposure to light from prolonged imaging. The experiments chosen in this study are representative of these cases. In Experiments 1, 3 and 6, the fluorescence videos show a complete pollen tube. Experiments 2, 4 and 5 have incomplete shapes in the fluorescence videos. These experiments also show that the cell can significantly change its shape over time. These changes would be difficult to track using shape models like Active Appearance Models [16] that will require user intervention to learn the new shape.

B. Performance Metrics

The proposed method and other segmentation algorithms are tested on experimental videos of growing pollen tubes.

TABLE IV
DESCRIPTION OF SYMBOLS IN FIGURE 7

Symbol	Description
A	Graph cut on the average of the last 5 fluorescence images from a given point
B	The proposed method using seam carving and feedback
C	Level set on the FEIF without feedback and seed points obtained from A above
D	Proposed method without feedback
E	Fusion + level set without feedback [25]
F	Level set on brightfield image with seed points from corresponding fluorescence image
G	Region growing on brightfield image with seed points from corresponding fluorescence image

The accuracy of each segmentation is computed using the following two metrics:

$$Error = \frac{1}{N_c}\sum_{i}^{N_c} d(C_i^0 G^0) + \frac{1}{N_g}\sum_{i}^{N_g} d(G_i^0, C^0) \quad (15a)$$

$$Dc = \frac{2|C \cap G|}{|C| + |G|} \quad (15b)$$

where C and G are binary masks for the segmented and groundtruth cell regions respectively. C^0 and G^0 are the contours of the segmented and groundtruth regions respectively. N_c and N_g are the lengths of the segmented and groundtruth contours, and $d(i, j)$ is the distance of point i to contour j. In equation (15a) the first term determines how close the segmented contour is to the groundtruth contour, and the second term measures how close the groundtruth contour is to the segmented contour. Thus a segmentation with a lower $Error$ value is more accurate than one with a higher value. The groundtruth masks (G) are obtained by manual drawing using ImageJ software [29]. The Dice coefficient (Dc) in equation (15b) measures the degree of overlap between the segmentation mask (C) and the groundtruth mask (G).

C. Experiments and Discussion

We perform segmentation using three algorithms (graph cut, seam carving and level set) under two conditions: with temporal considerations (feedback) and without temporal considerations (based on individual images). Table IV gives a description of the symbols used in Figure 7. Methods (B, C) use the proposed fusion and feedback, method (D) uses the proposed fusion without feedback, and methods (F, G) do not

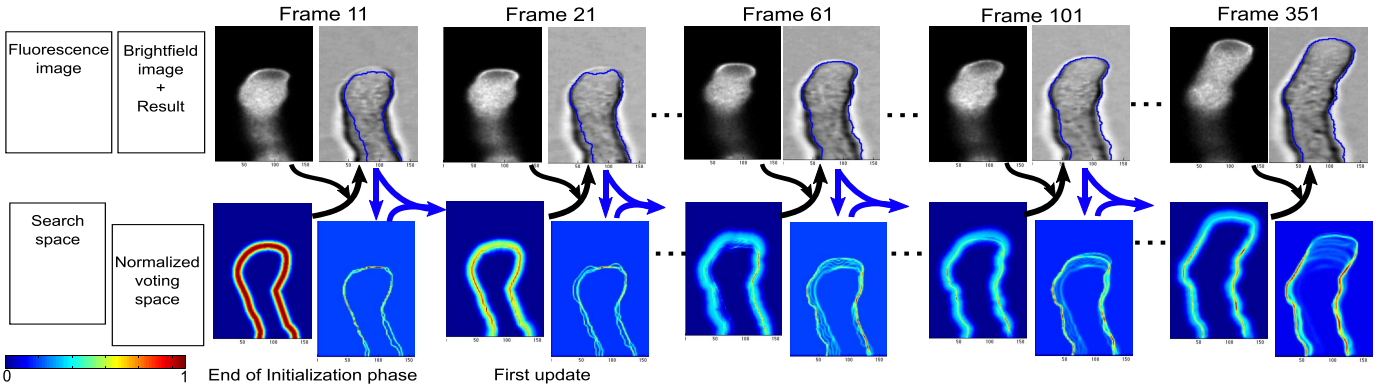

Fig. 6. Sample snapshots from Experiment #10 showing the initialization step, temporal updates of the search space, the normalized voting space and the segmentation result superimposed on the brightfield image for selected frames. At each frame, the updated search space (lower left) is used to obtain a segmentation (upper right). This segmentation is then used to update the voting space (lower right), and also used to update the search space for the next segmentation (lower left, next frame). The first search space update (Frame 21) is very similar to the initialized search space, since historical evidence is very low. Over time, the search space reflects historical segmentations, as evidenced by the high intensities in the cylindrical regions of the cell (Frame 351).

use fusion or feedback. The results of each method are shown in Figure 7. The following color codes and abbreviations are used: (green, GC) the result of graph cut; (blue) the proposed method; (red, cLS) the result of level set on the FEIF [25]; (black, LS) the result of level set [10] on a brightfield image; (yellow, RG) the result of region growing on a brightfield image.

The level set based methods (C, E, F) were set to terminate after 2000 iterations or when the change in the area of the segmented region is less than 5%. This change is determined after every 10^{th} iteration during the segmentation.

1) Sample Detailed Results: Figure 6 shows sample images and results from Experiment #10 for the proposed method (B) as well as updates for the search space. Frame 11 is the second frame used for initialization of the search space (Section II-B) that is used to compute the segmentation (up black arrow). The segmentation from this frame is used to update the voting space (down, blue arrow) as well as compute motion vectors that are used to update the search space for the next segmentation (right, blue arrow). Subsequent frames will use updated search spaces. The next analyzed image (Frame 21) uses the updated search space for shape detection and the resulting segmentation is used to update voting space and the next search space. Between Frames 21 and 61, the voting space shows shifting of the left boarder of the cell which indicates both growth and a correction of the local segmentation. The search space at Frame 101 shows two focus regions at the tip. The first is caused by immediate increase in motion vectors in the tip region and the second is due to previous segmentations in that region. This highlights the predictive attributes of the search space. The voting space in Frame 351 shows strong cell walls in the cylindrical region because previous segmentations have been in these particular regions. The tip is not as bright, indicating that these locations do not have as much historical evidence as the cylindrical regions. Also, some translation occurred during the growth sequence and this is reflected by the left-shift in the voting space at the base of the cell.

2) Results Without Feedback: (See Figure 7). Methods (D, E) do not use feedback, but use the proposed fusion technique. As such, each image is segmented independently and the initialization step (with or without shape estimation) is performed for every image. These methods generally show lower average error and higher Dice coefficients than their counterparts using temporal information (B, C) as demonstrated in Figure 7. However method (D) has a significantly lower processing time compared to method (E).

3) Results With Feedback: (See Figure 7). The result for the proposed fusion and feedback mechanism is indicated by method (B). Experiments 3, 6, 7 and 10 show significant difference between temporal and non-temporal based methods. In Experiments 3 and 6, the pollen tube is growing straight with some turning at the end. The fluorescence images show complete pollen tubes, but the brightfield images show diminishing edge contrast at the cell wall. The non-temporal methods are unaffected since their search spaces have a broader impact region compared to those generated through temporal updates (see search space image in Figure 6, Frame 11 vs. Frames 61, 101, 351). Even though the *Error* values for method B are higher than the other methods for these experiments, the Dice coefficients indicate that the segmented regions from method B are close to the groundtruth regions.

Experiments 7 and 10 (same video, but different sampling time) highlight the strength of the proposed feedback method. This video shows a growing pollen tube that turns and continues growing. The fluorescence images initially show an almost complete pollen tube, but gets worse over the course of the video and the brightfield images initially show some regions of weak edges at the tip region. Method (B) is able to adapt to the changing growth criteria while the other methods show significantly higher errors. Furthermore, the Dice coefficients also show that the area of the proposed method closely agrees with the groundtruth while the other methods do not. (Figure 7(b), column 7). Methods (C and D) fail because the shape estimation process expects a straight pollen tube due to the orientation of the fluorescence region during the course of the video. An incorrect region is selected as the search space which leads to poor results.

4) Processing Time: From Figure 8(b) the median processing time for the proposed method (1.491 seconds) is below

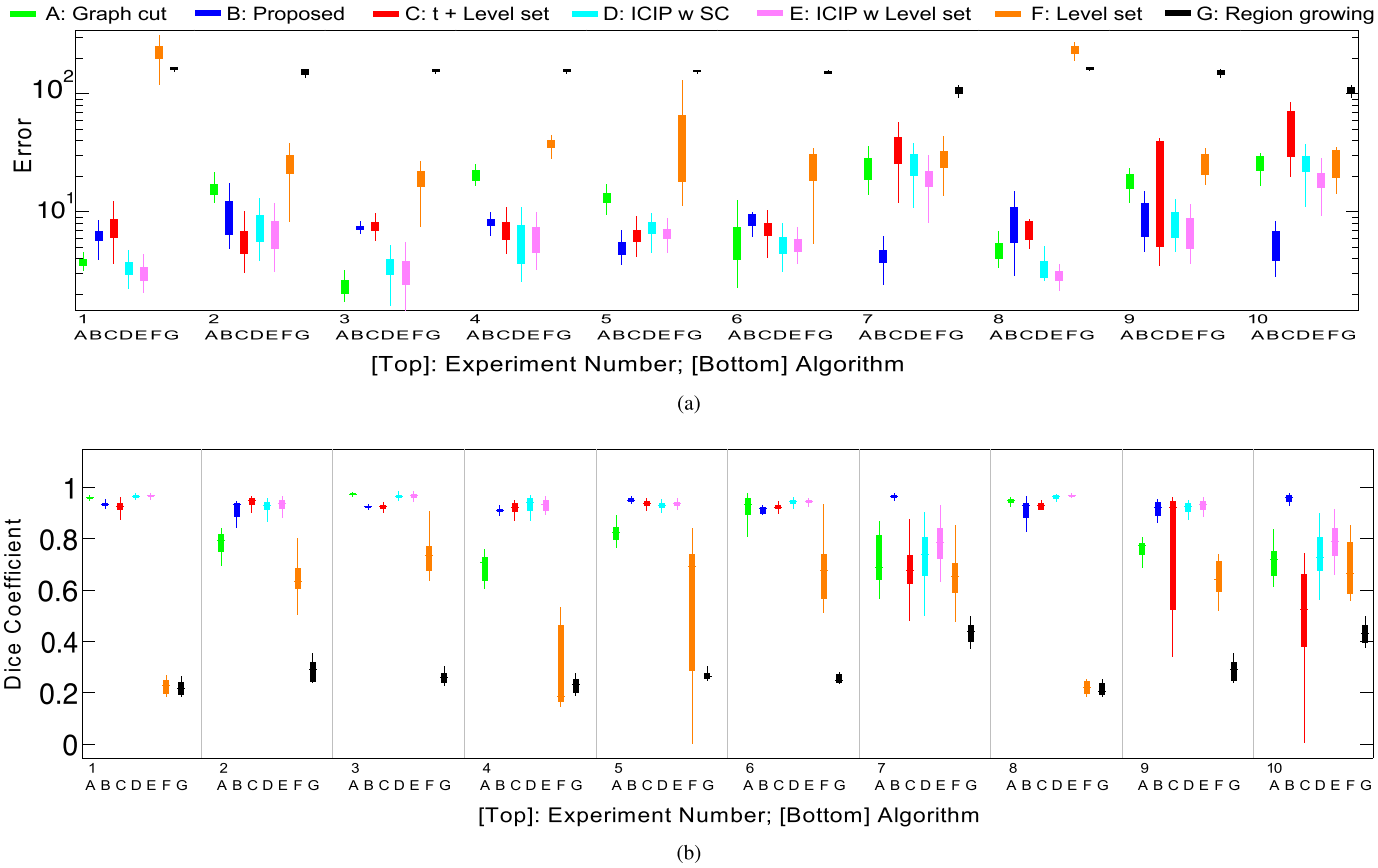

Fig. 7. Boxplot of experimental results. (A - G) are described at the beginning of Section III-C. (a) Box plot of the Error of each segmentation method for the experiments listed in Table III. (b) Dice coefficient for each segmentation methods per experiment.

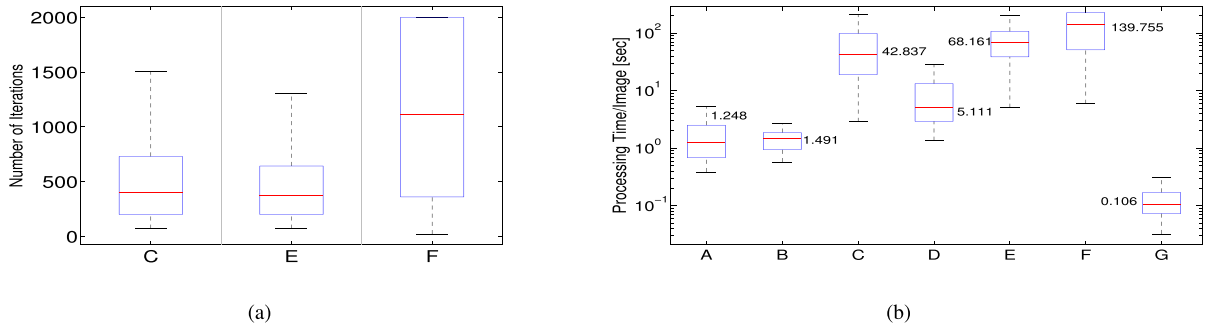

Fig. 8. Run time statistics for the segmentation methods. (a) Number of iterations required for the level set based methods (C, E, F). Median values are 400, 370 and 1115 iterations respectively. (b) Processing time per image for each algorithm over all experiments. Processing time was recorded on a 2.6GHz Intel core i5 laptop with 8GB of RAM.

the sampling time between consecutive imaging points in most of the experiments (3 sec). The standard deviation in the processing time is 0.4 seconds. Thus, this method can be applied during experimental imaging if the time between consecutive images is on average > 3 seconds. ($mean + 3\sigma$: Assuming a Gaussian distribution, a sampling time ≥ 2.691 seconds should account for at least 99% of processing times).

5) Sample Results: Figure 9 shows sample images from Experiment 1. In this sequence, the fluorescence video always has a complete cell shape, so the benefits of shape estimation are not evident. The brightfield images have a region of low gradient magnitudes in the tip region which causes the level set method (E) and region growing (F) methods to bleed out of the field of view (hence low Dice scores in Figure 7(b)). In particular, the failure of level set (E) and region growing (F) methods is significant since they do not benefit from the proposed fusion process. The proposed fusion method amplifies the weak gradients in the cell region and accurate segmentations are obtained. This situation occurs in Experiments 3, 6 and 8 where the fluorescence images show complete cell shapes.

Figure 10 shows sample results from Experiment #7. The fluorescence images show incomplete pollen tubes at the beginning of the video, which gets worse towards the end

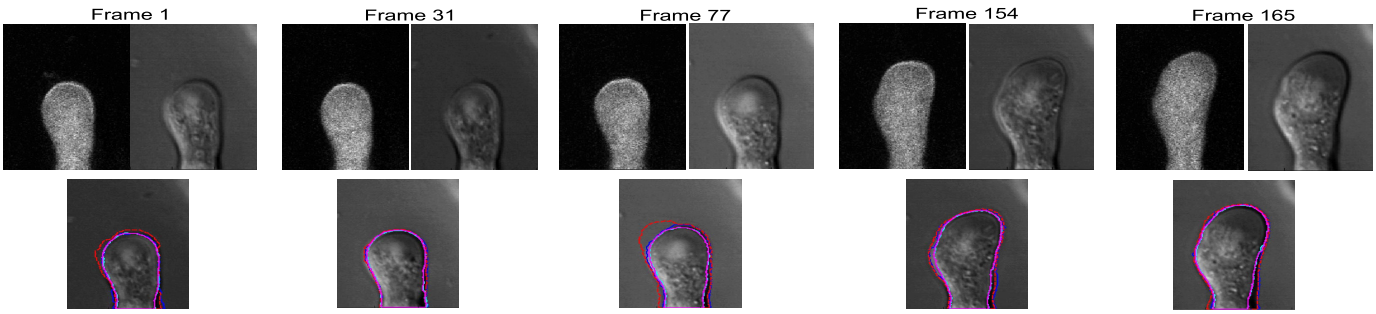

Fig. 9. Sample images from Experiment #7 showing the results of selected segmentation methods. Blue, method B (proposed); Red, method C; Cyan, method D; Magenta, method E. The color code matches Figure 7. The top row shows the pair of fluorescence and brightfield images for the indicated Frame number, and the bottom row shows the results of segmentation overlaid on the brightfield image. This represents the case were all fusion methods perform well.

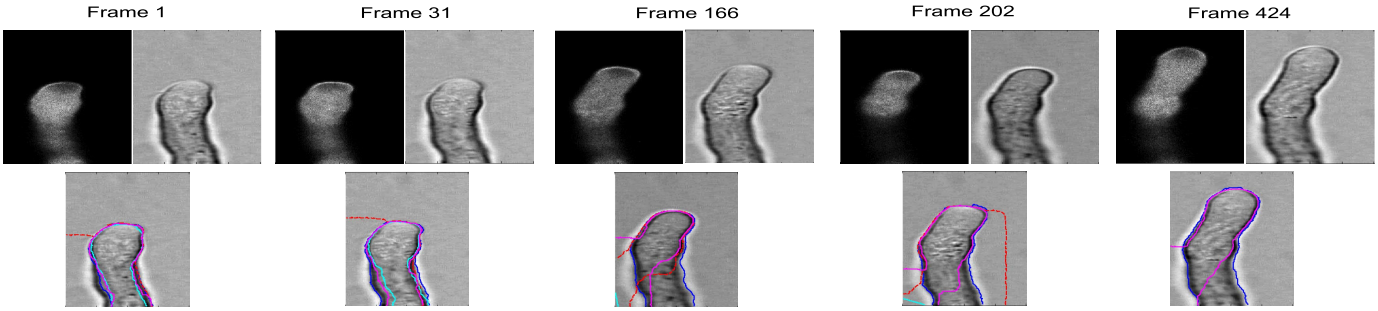

Fig. 10. Sample images from Experiment #7 showing the results of selected segmentation methods. The top row shows the pair of fluorescence and brightfield images for the indicated Frame number, and the bottom row shows the results of segmentation overlaid on the brightfield image. The color code matches Figure 7. Without temporal updates, most of the methods fail and bleed out. The proposed method with temporal update (Blue) is able to adjust to the growth behavior.

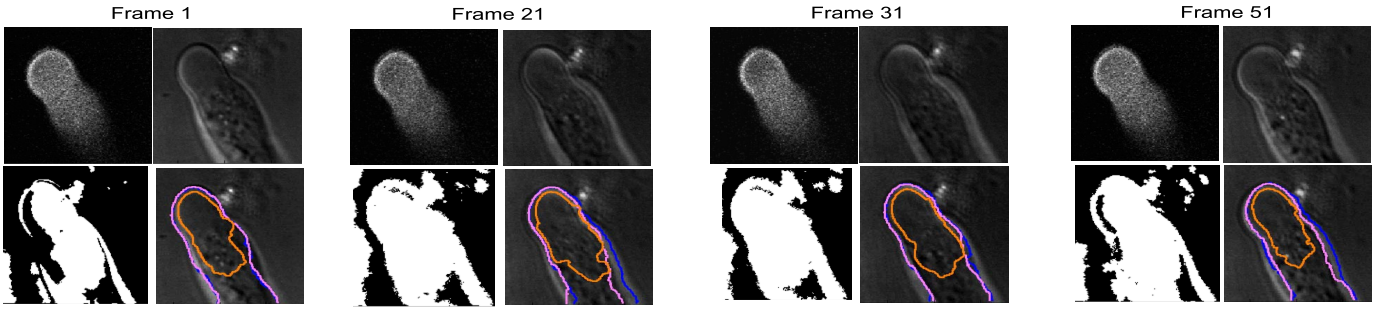

Fig. 11. Sample results from Experiment 2 showing results with (and without) the proposed fusion method (lower right) and the method by Tarabalka et al. [24] (lower left). Blue = proposed method, Magenta = [25], Orange = level set without fusion. The top row shows the corresponding fluorescence and brightfield images.

of the video. These produce incomplete segmentations for graph cut (method A) and are the reason for its low accuracy shown in Figure 7a. As such, the segmentation greatly benefits from the initial shape estimation and fusion processes. The brightfield images have strong edges. There is also a shift in the orientation of the cell during the course of the video with respect to the initial orientation. The proposed method B is able to adapt to the changing growth behavior while the other methods that rely entirely on shape estimation fail.

6) Comparison With Other Methods: Figure 11 shows sample images from the comparison of our fusion method with that of Tarabalka et al. [24], which is a graph cut based method with spatio-temporal constraints for shape growth or shrinkage. In this case, the constraints on the modalities is that the result for the fluorescence image should be included in that of the brightfield image. The intensities in the brightfield image corresponding to the foreground of the fluorescence image are similar to some background intensities, hence their inclusion in the final segmentation. As shown, our fusion method performs better because we are searching for pixels belonging to a boundary, not pixels belonging to a region.

IV. Conclusions

We described a dynamic method for fusing information from two different image modalities (fluorescence and brightfield) to achieve a unified segmentation of a growing pollen tube. The fusion leverages the localization abilities of fluorescence images and, together with shape estimation, enhances edges in the brightfield Image. Shape estimation is made possible by understanding the 2D shape of the pollen tube and creating a search space in the image plane within which the cell boundary will be found. Given that there are many such candidates for the cell boundary, the optimal boundary is found by minimizing the energy of the curve from one end of the shape

to the other. During video segmentation, we perform updates of the search space based on the current segmentation and motion vectors between the last two segmentations. With this update, the method no longer has to perform shape verification since this is relegated to the initialization phase. Experimental results show that the proposed boundary detection scheme with feedback is more robust to changing growth dynamics compared to those that only rely on shape estimation. Also, in the search for the best contour, seam carving-based segmentation was faster than level set based methods. In general, fusing both modalities lead to better segmentation than without fusion.

ACKNOWLEDGMENTS

The authors thank Dr. Zhenbiao Yang and Nan Luo for providing the videos used in this work.

REFERENCES

[1] C. M. Rounds, P. K. Hepler, and L. J. Winship, "The apical actin fringe contributes to localized cell wall deposition and polarized growth in the Lily pollen tube," *Plant Physiol.*, vol. 166, no. 1, pp. 139–151, Sep. 2014.

[2] Y. Qin and Z. Yang, "Rapid tip growth: Insights from pollen tubes," *Seminars Cell Develop. Biol.*, vol. 22, no. 8, pp. 816–824, Oct. 2011.

[3] A. L. Tambo, B. Bhanu, N. Ung, N. Thakoor, N. Luo, and Z. Yang, "Understanding pollen tube growth dynamics using the unscented Kalman filter," *Pattern Recognit. Lett.*, Aug. 2015. [Online]. Available: http://dx.doi.org/10.1016/j.patrec.2015.07.016

[4] J. Liu and P. J. Hussey, "Dissecting the regulation of pollen tube growth by modeling the interplay of hydrodynamics, cell wall and ion dynamics," *Frontiers Plant Sci.*, vol. 5, no. 8, p. 392, Aug. 2014.

[5] A. L. Tambo, B. Bhanu, N. Luo, G. Harlowt, and Z. Yang, "Integrated model for understanding pollen tube growth in video," in *Proc. 22nd Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2014, pp. 2727–2732.

[6] P. K. Hepler, C. M. Rounds, and L. J. Winship, "Control of cell wall extensibility during pollen tube growth," *Molecular Plant*, vol. 6, no. 4, pp. 998–1017, Jul. 2013.

[7] L. Steinhorst and J. Kudla, "Calcium—A central regulator of pollen germination and tube growth," *Biochim. Biophys. Acta-Molecular Cell Res.*, vol. 1833, no. 7, pp. 1573–1581, Jul. 2013.

[8] P. Fayant, O. Girlanda, Y. Chebli, C.-É. Aubin, I. Villemure, and A. Geitmann, "Finite element model of polar growth in pollen tubes," *Plant Cell*, vol. 22, no. 8, pp. 2579–2593, Aug. 2010.

[9] J. H. Kroeger, A. Geitmann, and M. Grant, "Model for calcium dependent oscillatory growth in pollen tubes," *J. Theor. Biol.*, vol. 253, no. 2, pp. 363–374, Jul. 2008.

[10] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3243–3254, Dec. 2010.

[11] H. Kimm, N. Abolhassani, and F. Lee, "Edge detection and linking pattern analysis using Markov chains," in *Proc. 16th Int. Conf. Comput. Sci. Eng.*, Dec. 2013, pp. 1146–1152.

[12] X. Ji, X. Zhang, and L. Zhang, "Sequential edge linking method for segmentation of remotely sensed imagery based on heuristic search," in *Proc. 21st Int. Conf. Geoinform.*, Jun. 2013, pp. 1–5.

[13] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[14] F. Yi and I. Moon, "Image segmentation: A survey of graph-cut methods," in *Proc. IEEE Int. Conf. Syst. Inform. (ICSAI)*, May 2012, pp. 1936–1941.

[15] W. Deng, W. Xiao, H. Deng, and J. Liu, "MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve," in *Proc. IEEE 3rd Int. Conf. Biomed. Eng. Inform.*, Oct. 2010, pp. 393–396.

[16] T. F. Cootes and C. J. Taylor, "An algorithm for tuning an active appearance model to new data," in *Proc. Brit. Mach. Vis. Conf.*, 2006, p. 94.1. [Online]. Available: http://www.bmva.org/bmvc/2006/papers/040.html.

[17] X. Gao, Y. Su, X. Li, and D. Tao, "A review of active appearance models," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 40, no. 2, pp. 145–158, Mar. 2010. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5373937.

[18] J. Han and B. Bhanu, "Detecting moving humans using color and infrared video," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Jul./Aug. 2003, pp. 228–233.

[19] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 7, pp. 903–921, Jul. 2004.

[20] X. Zou and B. Bhanu, "Tracking humans using multi-modal fusion," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.-Workshops (CVPR)*, Jun. 2005, p. 4.

[21] B. Parvin, Q. Yang, J. Han, H. Chang, B. Rydberg, and M. H. Barcellos-Hoff, "Iterative voting for inference of structural saliency and characterization of subcellular events," *IEEE Trans. Med. Imag.*, vol. 16, no. 3, pp. 615–623, Mar. 2007.

[22] A. J. Asman and B. A. Landman, "Robust statistical label fusion through consensus level, labeler accuracy, and truth estimation (COLLATE)," *IEEE Trans. Med. Imag.*, vol. 30, no. 10, pp. 1779–1794, Oct. 2011.

[23] X. Liu, A. Montillo, E. T. Tan, and J. F. Schenck, "iSTAPLE: Improved label fusion for segmentation by combining STAPLE with image intensity," *Proc. SPIE 8669, Med. Imag., Image Process.*, pp. 86692O-1–86692O–6, Mar. 2013, doi: 10.1117/12.2006447.

[24] Y. Tarabalka, G. Charpiat, L. Brucker, and B. H. Menze, "Spatio-temporal video segmentation with shape growth or shrinkage constraint," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3829–3840, Sep. 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25020092.

[25] A. L. Tambo and B. Bhanu, "Dynamic bi-modal fusion of images for the segmentation of pollen tubes in video," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 148–152.

[26] A. Shamir and S. Avidan, "Seam carving for media retargeting," *Commun. ACM*, vol. 3, no. 1, pp. 77–85, 2009.

[27] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.

[28] B. Bhanu and W. Burger, "Approximation of displacement fields using wavefront region growing," *Comput. Vis., Graph., Image Process.*, vol. 41, no. 3, pp. 306–322, 1988.

[29] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature*, vol. 9, no. 7, pp. 671–675, Jul. 2012.

Asongu L. Tambo (S'11) received the B.A. degree in computer science and physics from the University of Minnesota at Morris, in 2006, and the M.S. degree in electrical engineering from California State University at Northridge, in 2009. He is currently pursuing the Ph.D. degree in electrical engineering with the Center for Research in Intelligent Systems, University of California, Riverside. His research interests include image processing, pattern recognition, computer vision, machine learning, and modeling biological dynamics. His current research focuses on the dynamics of pollen tube growth.

Bir Bhanu (F'95) received the S.M. and E.E. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, the Ph.D. degree in electrical engineering from the Image Processing Institute, University of Southern California, and the M.B.A. degree from the University of California, Irvine. He is currently the Distinguished Professor of Electrical and Computer Engineering and the Cooperative Professor of Computer Science and Engineering, Mechanical Engineering and Bioengineering, the Director of the Center for Research in Intelligent Systems, and the Visualization and Intelligent Systems Laboratory with the University of California, Riverside (UCR). He also serves as the Director of NSF IGERT Program on Video Bioinformatics and the Interim Chair of the Department of Bioengineering. Prior to that he was a Senior Honeywell Fellow at Honeywell Inc. He was the First Founding Faculty Member with the Bourns College of Engineering and served as the Chair of Electrical Engineering at UCR. His research interests are computer vision, pattern recognition and data mining, machine learning, artificial intelligence, image processing, image and video database, biological, medical, military, and intelligence applications. He has been the Principal Investigator of programs from NSF, DARPA, NASA, AFOSR, ONR, ARO, and other agencies and industries. He is a fellow of AAAS, IAPR, AIMBE, and SPIE.